



①⑨ **BUNDESREPUBLIK
DEUTSCHLAND**



**DEUTSCHES
PATENT- UND
MARKENAMT**

⑫ **Offenlegungsschrift**
⑩ **DE 102 49 060 A 1**

⑤① Int. Cl.⁷:
G 10 L 15/22
G 06 F 3/00

②① Aktenzeichen: 102 49 060.0
②② Anmeldetag: 22. 10. 2002
④③ Offenlegungstag: 27. 11. 2003

DE 102 49 060 A 1

⑥⑥ Innere Priorität:
102 21 490. 5 14. 05. 2002

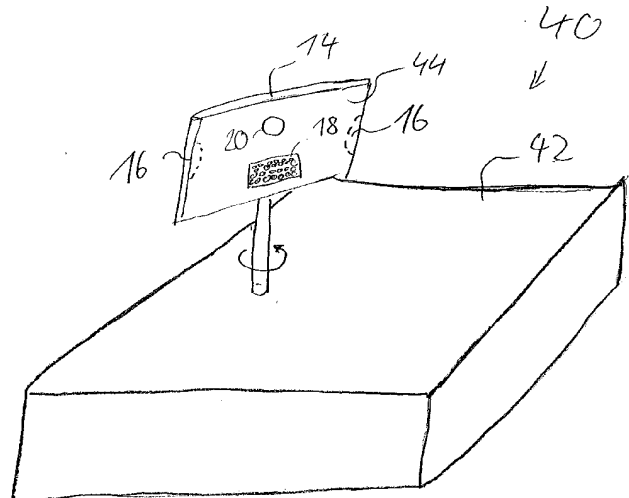
⑦① Anmelder:
Philips Intellectual Property & Standards GmbH,
20099 Hamburg, DE

⑦② Erfinder:
Oerder, Martin, Dr., 52134 Herzogenrath, DE

Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen

⑤④ Dialogsteuerung für elektrisches Gerät

⑤⑦ Eine Vorrichtung mit Mitteln zur Aufnahme und Erkennung von Sprachsignalen und ein Verfahren zur Steuerung eines elektrischen Geräts werden vorgestellt. Die Vorrichtung weist ein Verkörperungselement 14 auf, das mechanisch bewegt werden kann. Die Position eines Benutzers wird ermittelt und das Verkörperungselement 14, das bspw. die Darstellung eines menschlichen Gesichts umfassen kann, wird so bewegt, dass seine Vorderseite 44 in Richtung der Position des Benutzers weist. Am Verkörperungselement 14 können Mikrofone 16, Lautsprecher 18 und/oder eine Kamera 20 vorgesehen sein. Der Benutzer kann einen Sprach-Dialog mit der Vorrichtung führen, wobei sich ihm das Gerät in Form des Verkörperungselements 14 darstellt. Entsprechend der Spracheingaben des Benutzers kann ein elektrisches Gerät angesteuert werden. Ebenso ist ein Dialog des Benutzers mit dem Verkörperungselement möglich, mit dem Ziel, den Benutzer zu unterrichten.



DE 102 49 060 A 1

[0001] Die Erfindung betrifft eine Vorrichtung mit Mitteln zur Aufnahme und Erkennung von Sprachsignalen sowie ein Verfahren zur Kommunikation eines Benutzers mit einem elektronischen Gerät.

[0002] Es sind Spracherkennungsmittel bekannt, mit denen aufgenommenen akustischen Sprachsignalen das entsprechende Wort bzw. eine entsprechende Wortfolge zugeordnet werden kann. Spracherkennungssysteme werden – häufig als Dialogsysteme in Kombination mit Sprachsynthese – zur Steuerung elektrischer Geräte eingesetzt. Hierbei kann ein Dialog mit dem Benutzer als alleinige Schnittstelle für die Bedienung des elektrischen Geräts genutzt werden. Ebenso ist es möglich, dass die Sprachein- und ggfs. auch -ausgabe als eines von mehreren Kommunikationsmitteln genutzt wird.

[0003] In der US-A-6,118,888 ist eine Steuervorrichtung und ein -verfahren zur Steuerung eines elektrischen Geräts, bspw. eines Computers oder eines Geräts der Unterhaltungselektronik beschrieben. Für die Steuerung des Geräts stehen dem Benutzer eine Anzahl von Eingabemöglichkeiten zur Verfügung. Dies umfasst mechanische Eingabemöglichkeiten wie Tastatur oder Maus ebenso wie Spracherkennung. Zusätzlich verfügt das Steuergerät über eine Kamera, mit der Gestik und Mimik des Benutzers aufgenommen werden können, die als weitere Eingabesignale verarbeitet werden. Die Kommunikation mit dem Benutzer erfolgt in Form eines Dialogs, wobei dem System auch eine Anzahl von Wegen zur Verfügung stehen, um dem Benutzer Informationen zu übermitteln. Dies umfasst Sprachsynthese und -ausgabe. Insbesondere umfasst dies auch eine anthropomorphe Repräsentation, bspw. eine Darstellung eines Menschen, eines menschlichen Gesichts oder eines Tieres. Diese Repräsentation wird dem Benutzer in Form einer Computergraphik auf einem Bildschirm gezeigt.

[0004] Während Dialogsysteme in speziellen Anwendungen – bspw. in Telefon-Auskunftssystemen – heute bereits eingesetzt werden, ist in anderen Bereichen – bspw. Steuerung von elektrischen Geräten im Haushalt, Unterhaltungselektronik – die Akzeptanz noch gering.

[0005] Es ist Aufgabe der Erfindung, eine Vorrichtung mit Mitteln zur Aufnahme zur Erkennung von Sprachsignalen sowie ein Bedienverfahren für ein elektronisches Gerät vorzuschlagen, bei dem einem Benutzer die Bedienung mittels Sprachsteuerung besonders leicht fällt.

[0006] Diese Aufgabe wird gelöst durch eine Vorrichtung nach Anspruch 1 und ein Verfahren nach Anspruch 11. Abhängige Ansprüche beziehen sich auf vorteilhafte Ausführungsformen der Erfindung.

[0007] Die erfindungsgemäße Vorrichtung weist ein mechanisch bewegbares Verkörperungselement auf. Hierbei handelt es sich um einen Teil der Vorrichtung, der für den Benutzer als Verkörperung eines Dialogpartners dient. Die konkrete Ausgestaltung eines solchen Verkörperungselements kann sehr verschieden ausfallen. Bspw. kann es sich um einen Gehäuseteil handeln, der gegenüber einem Standgehäuse eines elektrischen Geräts motorisch beweglich angeordnet ist. Maßgeblich ist, dass das Verkörperungselement eine für den Benutzer als solche erkennbare Vorderseite aufweist. Ist diese Vorderseite dem Benutzer zugewandt, so wird ihm hierdurch der Eindruck vermittelt, dass das Gerät "aufmerksam" ist, d. h. Sprachkommandos entgegennehmen kann.

[0008] Erfindungsgemäß verfügt die Vorrichtung über Mittel zur Bestimmung der Position eines Benutzers. Dies kann bspw. über akustische oder optische Sensoren erfolgen. Die Bewegungsmittel des Verkörperungselements wer-

den so angesteuert, dass die Vorderseite des Verkörperungselements in Richtung der Position des Benutzers weist. So hat der Benutzer stets den Eindruck, dass die Vorrichtung bereit ist, ihm "zuzuhören".

[0009] Gemäß einer Weiterbildung der Erfindung umfasst das Verkörperungselement eine anthropomorphe Darstellung. Hierbei kann es sich um eine Darstellung eines Menschen oder eines Tieres, aber auch um eine Phantasiefigur (bspw. ein Roboter) handeln. Bevorzugt wird eine Darstellung eines menschlichen Gesichts. Hierbei kann es sich um eine realistische oder auch lediglich symbolische Darstellung handeln, bei der bspw. lediglich die Umrisse sowie Augen, Nase und Mund dargestellt sind.

[0010] Bevorzugt wird, dass die Vorrichtung auch Mittel zur Ausgabe von Sprachsignalen umfasst. Für die Steuerung eines elektronischen Geräts ist zwar vor allem die Spracherkennung maßgeblich. Rückmeldungen, Bestätigungen, Nachfragen etc. können jedoch mit Sprachausgabemitteln realisiert werden. Dies kann die Wiedergabe vorgespeicherter Sprachsignalen ebenso wie echte Sprachsynthese umfassen. Mit Sprachausgabemitteln kann eine vollständige Dialogsteuerung realisiert sein. Es können auch Dialoge mit dem Benutzer geführt werden, um diesen zu unterhalten.

[0011] Gemäß einer Weiterbildung der Erfindung umfasst die Vorrichtung eine Anzahl von Mikrofonen und/oder mindestens eine Kamera. Die Aufnahme von Sprachsignalen kann bereits mit einem einzigen Mikrofon erreicht werden. Bei der Verwendung von mehreren Mikrofonen kann aber einerseits eine Richtcharakteristik erreicht werden. Andererseits ist durch Empfang des Sprachsignals eines Benutzers mit mehreren Mikrofonen auch eine Ortung der Position des Benutzers möglich. Mit einer Kamera kann die Umgebung des Geräts beobachtet werden. Durch entsprechende Bildverarbeitung kann aus dem aufgenommenen Bild ebenfalls die Position des Benutzers ermittelt werden. Die Mikrofone, die Kamera und/oder Lautsprecher zur Ausgabe von Sprachsignalen können am mechanisch beweglichen Verkörperungselement angebracht sein. Bspw. bei einem Verkörperungselement, dass einem menschlichen Kopf nachgebildet ist, können im Bereich der Augen zwei Kameras, anstelle des Mundes ein Lautsprecher und im Bereich der Ohren zwei Mikrofone angeordnet sein.

[0012] Bevorzugt wird, dass Mittel zur Identifikation eines Benutzers vorgesehen sind. Dies kann bspw. durch Auswertung eines aufgenommenen Bildsignals (Gesichtserkennung) oder durch Auswertung des aufgenommenen akustischen Signals (Sprechererkennung) erreicht werden. Hiermit kann die Vorrichtung aus einer Anzahl von Personen, die sich in der Umgebung aufhalten, den jeweils aktuellen Benutzer ermitteln und das Verkörperungselement auf diesen Benutzer ausrichten.

[0013] Für die Bewegungsmittel zur mechanischen Bewegung des Verkörperungselementes gibt es die verschiedensten Möglichkeiten. Bspw. können diese Mittel Elektromotoren oder hydraulische Stellmittel umfassen. Das Verkörperungselement kann auch durch die Bewegungsmittel fortbewegt werden. Bevorzugt wird jedoch, dass das Verkörperungselement gegenüber einem feststehenden Teil lediglich schwenkbar ist. Hierbei sind bspw. Schwenkbewegungen um eine horizontale und/oder vertikale Achse möglich.

[0014] Die erfindungsgemäße Vorrichtung kann Teil eines elektrischen Geräts sein. Hierbei kommen bspw. Geräte der Unterhaltungselektronik (z. B. TV, Abspielgeräte für Audio- und/oder Video etc.) in Frage. Die Vorrichtung stellt in diesem Fall die Benutzerschnittstelle für das Gerät dar. Zusätzlich kann das Gerät auch über andere Bedienungsmittel (Tastatur etc.) verfügen. Alternativ kann die erfindungsgemäße Vorrichtung auch ein eigenständiges Gerät sein, das als

Steuervorrichtung zur Steuerung eines oder mehrerer separater elektrischer Geräte dient. In diesem Fall verfügen die zu steuernden Geräte über einen elektrischen Steueranschluss (z. B. drahtlos oder über einen geeigneten Steuerbus), über den die Vorrichtung die Geräte entsprechend der vom Benutzer entgegengenommenen Sprachkommandos steuert.

[0015] Insbesondere kann die erfindungsgemäße Vorrichtung für den Benutzer als Schnittstelle eines Systems zur Datenablage und/oder Abfrage dienen. Hierfür sind in der Vorrichtung interne Datenspeicher vorhanden oder die Vorrichtung ist – bspw. über ein Computer-Netzwerk oder das Internet – an externe Datenspeicher angebunden. Im Dialog kann der Benutzer Daten ablegen (z. B. Telefonnummern speichern, Memos ablegen etc.) oder Daten abfragen (beispielsweise Uhrzeit, Nachrichten, das aktuelle Fernsehprogramm etc.).

[0016] Darüber hinaus können die Dialoge mit dem Benutzer auch genutzt werden, um Parameter der Vorrichtung selbst einzustellen und deren Konfiguration zu ändern.

[0017] Wenn ein Lautsprecher zur Ausgabe und ein Mikrofon zur Aufnahme von akustischen Signalen vorgesehen sind, so kann eine Signalverarbeitung mit Störunterdrückung vorgesehen sein, d. h. eine Verarbeitung der aufgenommenen akustischen Signale so, dass Anteile des vom Lautsprecher ausgehenden akustischen Signals unterdrückt werden. Dies ist insbesondere vorteilhaft, wenn Lautsprecher und Mikrofon in räumlicher Nähe, bspw. am Verkörperungselement angebracht sind.

[0018] Zusätzlich zur oben angegebenen Verwendung der Vorrichtung zur Steuerung eines elektrischen Geräts kann diese auch zur Führung eines Dialogs mit dem Benutzer verwendet werden, der anderen Zwecken, bspw. der Information, Unterhaltung oder Unterrichtung des Benutzers dient. Gemäß einer Weiterbildung der Erfindung sind Dialogmittel vorgesehen, mit denen ein Dialog zur Unterrichtung des Benutzers geführt werden kann. Der Dialog läuft dann bevorzugt so ab, dass dem Benutzer Aufgaben gestellt und seine Antworten aufgenommen werden. Bei den Aufgaben kann es sich um komplexe Fragestellungen handeln, bevorzugt wird aber ein Abfragen von kurzen Lerngegenständen wie bspw. Vokabeln einer fremden Sprache, bei denen Aufgabe (bspw.: Definition eines Wortes) und Antwort (bspw.: Wort in der Fremdsprache) relativ kurz sind. Der Dialog wird vom Benutzer mit dem Verkörperungselement geführt und kann visuell und/oder per Audio erfolgen.

[0019] Als eine Möglichkeit einer effektiven Lernmethode wird vorgeschlagen, dass ein Satz von Lerngegenständen (bspw.: Vokabeln einer Fremdsprache) gespeichert sind, wobei für jeden Lerngegenstand mindestens gespeichert ist eine Aufgabe (bspw.: Definition), eine Lösung (bspw.: Vokabel) und eine Maßzahl für die Zeitdauer seit der letzten Stellung der Aufgabe an den Benutzer bzw. der korrekten Lösung der Aufgabe durch diesen. Während des Dialogs werden nun nacheinander Lerngegenstände ausgewählt und abgefragt, indem die Aufgabe dem Benutzer gestellt und die Antwort des Benutzers mit der gespeicherten Lösung verglichen wird. Die Auswahl des jeweils abzufragenden Lerngegenstands berücksichtigt hierbei die gespeicherte Maßzahl, d. h. die seit der letzten Abfrage des Gegenstands vergangene Zeit. Dies kann bspw. über ein geeignetes Lernmodell mit einer angenommenen oder ermittelten Vergessensrate erfolgen. Zusätzlich kann jeder Lerngegenstand auch mit einer Relevanz-Maßzahl bewertet sein, wobei die Relevanz-Maßzahl zusätzlich zur Zeitdauer-Maßzahl bei der Auswahl berücksichtigt wird.

[0020] Nachfolgend werden Ausführungsformen der Erfindung anhand von Zeichnungen näher beschrieben. Hier-

bei zeigen:

[0021] Fig. 1 ein Blockdiagramm von Elementen einer Steuervorrichtung;

[0022] Fig. 2 eine perspektivische Darstellung eines elektrischen Geräts mit einer Steuervorrichtung.

[0023] Fig. 1 zeigt in einem Blockdiagramm eine Steuervorrichtung **10** und ein hierdurch angesteuertes Gerät **12**. Die Steuervorrichtung **10** tritt für den Benutzer durch ein Verkörperungselement **14** in Erscheinung. Am Verkörperungselement **14** sind ein Mikrofon **16**, ein Lautsprecher **18** und ein Positionssensor für die Position eines Benutzers, hier in Form einer Kamera **20** angebracht. Gemeinsam bilden diese Elemente eine mechanische Einheit **22**. Das Verkörperungselement **14** und damit die mechanische Einheit **22** wird durch einen Motor **24** um eine vertikale Achse verschwenkt. Eine zentrale Steuereinheit **26** steuert über eine Ansteuerschaltung **28** den Motor **24** an. Das Verkörperungselement **14** ist eine eigenständige mechanische Einheit. Es weist eine für den Benutzer als solche erkennbare Vorderseite auf. Mikrofon **16**, Lautsprecher **18** und Kamera **20** sind am Verkörperungselement **14** in Richtung dieser Vorderseite angebracht.

[0024] Das Mikrofon **16** liefert ein akustisches Signal. Dieses wird von einem Aufnahmesystem **30** aufgenommen und durch eine Spracherkennungseinheit **32** bearbeitet. Das Ergebnis der Spracherkennung, d. h. die dem aufgenommenen akustischen Signal zugeordnete Wortfolge wird an die zentrale Steuereinheit **26** geleitet.

[0025] Die zentrale Steuereinheit **26** steuert auch eine Sprachsyntheseeinheit **34** an, die ein synthetisches Sprachsignal über eine Tonerzeugungseinheit **36** und den Lautsprecher **18** ausgibt.

[0026] Das von der Kamera **20** aufgenommene Bild wird von der Bildverarbeitungseinheit **38** verarbeitet. Die Bildverarbeitungseinheit **38** ermittelt aus dem von der Kamera **20** gelieferten Bildsignal den Standort eines Benutzers. Die Standortinformation wird an die zentrale Steuereinheit **26** übermittelt.

[0027] Die mechanische Einheit **22** dient als Benutzerschnittstelle, über die die zentrale Steuereinheit **26** Eingaben des Benutzers entgegennimmt (Mikrofon **16**, Spracherkennungseinheit **32**) und Rückmeldungen an den Benutzer gibt (Sprachsyntheseeinheit **34**, Lautsprecher **18**). Die Steuereinheit **10** dient hierbei zur Steuerung eines elektrischen Geräts **12**, bspw. eines Geräts aus dem Bereich der Unterhaltungselektronik.

[0028] Die Darstellung der Steuervorrichtung **10** in Fig. 1 zeigt lediglich symbolisch deren Funktionseinheiten. Die verschiedenen Einheiten, bspw. zentrale Steuereinheit **26**, Spracherkennungseinheit **32**, Bildverarbeitungseinheit **38** können in einer konkreten Umsetzung als separate Baugruppen vorhanden sein. Ebenso ist aber eine rein softwaremäßige Implementation dieser Einheiten denkbar, wobei die Funktionalität von mehreren oder allen dieser Einheiten durch ein auf einer Zentraleinheit ablaufendes Programm realisiert wird.

[0029] Ebenso ist es nicht zwingend, dass sich diese Einheiten in räumlicher Nähe zueinander oder zur mechanischen Einheit **22** befinden. Die mechanische Einheit **22**, d. h. das Verkörperungselement **14** sowie auch die bevorzugt, aber nicht zwingend daran angeordneten Einheiten Mikrofon **16**, Lautsprecher **18** und Sensor **20**, können entfernt vom Rest der Steuervorrichtung **10** separat angeordnet sein und lediglich – leitungsgebunden oder auch drahtlos – in Signalverbindung hiermit stehen.

[0030] Im Betrieb der Steuervorrichtung **10** wird ständig ermittelt, ob sich ein Benutzer in der Umgebung aufhält. Die Position des Benutzers wird ermittelt. Die zentrale Steuer-

einheit **26** steuert den Motor **24** so an, dass die Vorderseite des Verkörperungselements **10** in Richtung des Benutzers weist.

[0031] Die Bildverarbeitungseinheit **38** umfasst auch eine Gesichtserkennung. Liefert die Kamera **20** ein Bild mit mehreren Personen, so wird mit Hilfe der Gesichtserkennung ermittelt, welche der Personen der – dem System bekannte – Benutzer ist. Das Verkörperungselement **14** wird in Richtung dieses Benutzers ausgerichtet. Sind mehrere Mikrofone vorhanden, so können die Signale dieser Mikrofone so verarbeitet werden, dass eine Richtcharakteristik in Richtung der bekannten Position des Benutzer entsteht.

[0032] Die Bildverarbeitungseinheit **38** kann zusätzlich so ausgestaltet sein, dass sie die durch die Kamera **20** aufgenommene Szene in der Umgebung der mechanischen Einheit **22** "verstellt". Die jeweilige Szene kann dann einer von einer Anzahl vordefinierten Zustände zugeordnet werden. Bspw. ist der zentralen Steuereinheit **26** auf diese Weise bekannt, ob eine oder mehrere Personen im Raum sind. Die Einheit kann auch bspw. das Verhalten des Benutzers erkennen und zuordnen, d. h. z. B. ob der Benutzer in Richtung der mechanischen Einheit **22** blickt oder ob er sich im Gespräch mit einer anderen Person befindet. Durch Auswertung der so erkannten Zustände kann die Erkennungsleistung deutlich verbessert werden. Bspw. kann vermieden werden, dass Teile eines Gesprächs zwischen zwei Personen irrtümlich als Sprachkommandos interpretiert werden.

[0033] Die zentrale Steuereinheit ermittelt in einem Dialog mit dem Benutzer Eingaben und steuert das Gerät **12** entsprechend an. Ein solcher Dialog zur Steuerung der Lautstärke eines Audio-Wiedergabegeräts **12** kann bspw. wie folgt ablaufen:

- Der Benutzer wendet sich von seinem jeweiligen Standort dem Verkörperungselement **14** zu. Das Verkörperungselement **14** ist durch den Motor **24** stets so gerichtet, dass seine Vorderseite dem Benutzer zugewandt ist. Hierfür wird die Ansteuerschaltung **28** von der zentralen Steuereinheit **26** des Geräts **10** entsprechend der ermittelten Position des Benutzers angesteuert.
- Der Benutzer gibt ein Sprachkommando, bspw. "Lautstärke Fernseher". Das Sprachkommando wird vom Mikrofon **16** aufgenommen und durch die Spracherkennungseinheit **32** erkannt.
- Die zentrale Steuereinheit **26** reagiert mit einer Rückfrage: "Lauter oder leiser?", die über die Sprachsyntheseinheit **34** vom Lautsprecher **18** ausgegeben wird.
- Der Benutzer gibt das Sprachkommando "Leiser". Nach Erkennung des Sprachsignals steuert die zentrale Steuereinheit **26** das Gerät **12** so an, dass die Lautstärke verringert wird.

[0034] Fig. 2 zeigt eine perspektivische Darstellung eines elektronischen Geräts **40** mit einer integrierten Steuervorrichtung. Von der Steuervorrichtung **10** ist hier nur das Verkörperungselement **14** zu sehen, das gegenüber einem Standgehäuse **42** des Geräts **40** um eine vertikale Achse verschwenkbar ist. In diesem Beispiel ist das Verkörperungselement von flacher, rechteckiger Form. Auf der Vorderseite **44** befindet sich das Objektiv der Kamera **20** sowie der Lautsprecher **18**. Seitlich sind zwei Mikrofone **16** angeordnet. Durch einen Motor (nicht dargestellt) wird die mechanische Einheit **22** stets so gedreht, dass die Vorderseite in Richtung des Benutzers zeigt.

[0035] In einer Ausführungsform (nicht dargestellt) wird die Vorrichtung **10** aus Fig. 1 nicht zur Steuerung des Geräts

12, sondern zur Führung eines Dialogs mit dem Ziel der Unterrichtung eines Benutzers verwendet. Die zentrale Steuereinheit **26** führt hierbei ein Lernprogramm aus, mit dessen Hilfe der Benutzer eine Fremdsprache erlernen kann.

[0036] In einem Speicher ist ein Satz von Lerngegenständen abgespeichert. Hier handelt es sich um einzelne Datensätze, von denen jeder die Definition eines Wortes, das entsprechende Wort in der Fremdsprache, eine Bewertungsmaßzahl für die Relevanz des Wortes (Auftretenshäufigkeit des Wortes in der Sprache) und eine Zeit-Maßzahl für die Dauer der seit der letzten Abfrage des Datensatzes vergangene Zeit aufweist.

[0037] Eine Lerneinheit im Dialog läuft nun ab, indem nacheinander Datensätze ausgewählt und abgefragt werden. Hierbei wird dem Benutzer eine Aufgabe gestellt, d. h. die in dem Datensatz gespeicherte Definition optisch angezeigt oder akustisch ausgegeben. Die Antwort des Benutzers, bspw. eingegeben über eine Tastatur, bevorzugt aufgenommen über das Mikrofon **16** und die automatische Spracherkennung **32** wird aufgenommen und mit der gespeicherten Lösung (Vokabel) verglichen. Der Benutzer wird darüber informiert, ob die Lösung als richtig erkannt wurde. Für den Fall fehlerhafter Antworten kann dem Benutzer die richtige Lösung mitgeteilt oder zuvor ein- oder mehrmals Gelegenheit zur weiteren Antwort gegeben werden. Nach derartiger Abarbeitung des Datensatzes wird die abgespeicherte Maßzahl für die Zeitdauer seit der letzten Abfrage aktualisiert, d. h. auf Null gesetzt.

[0038] Danach erfolgt die Auswahl und Abfrage eines weiteren Datensatzes usw..

[0039] Die Auswahl der jeweils abzufragenden Datensätze erfolgt anhand eines Gedächtnismodells. Ein einfaches Gedächtnismodell wäre die Formel

$$P(k) = \exp(-t(k) \cdot r(c(k))),$$

wobei $P(k)$ die Wahrscheinlichkeit bezeichnet, dass der Lerngegenstand k bekannt ist, \exp die Exponentialfunktion bezeichnet, $t(k)$ die Zeit bezeichnet, seit der Gegenstand zuletzt abgefragt wurde, $c(k)$ die Lern-Klasse des Gegenstandes ist und $r(c(k))$ die lernklassen-spezifische Vergessensrate. Für t kann die Uhrzeit verwendet werden. Ebenso kann die Zeit t in Lernschritten gerechnet werden. Lern-Klassen können auf verschiedene geeignete Arten definiert werden.

Ein mögliches Modell ist es, für jedes $N > 0$ alle Gegenstände die N mal richtig beantwortet wurden einer entsprechenden Klasse zuzuordnen. Für die Vergessensrate kann ein geeigneter fester Wert angenommen werden, oder es kann ein geeigneter Startwert ausgewählt und bspw. durch einen Gradienten-Algorithmus adaptiert werden.

[0040] Ziel der Unterrichtung ist eine Maximierung eines Wissensmaßes. Dieses Wissensmaß wird definiert als der Anteil der dem Benutzer bekannten Lerngegenstände des Satzes, gewichtet mit der Relevanz-Maßzahl. Da die Abfrage eines Gegenstandes k die Wahrscheinlichkeit $P(k)$ auf Eins bringt, wird zur Optimierung der Wissens-Maßzahl vorgeschlagen, in jedem Schritt den Gegenstand mit der geringsten Wissenswahrscheinlichkeit $P(k)$, ggfs. gewichtet mit der Relevanz-Maßzahl $U(k)$, $U(k) \cdot 1-P(k)$ abzufragen. Mit Hilfe des Modells kann die Wissens-Maßzahl nach jedem Schritt berechnet und dem Benutzer angezeigt werden. Die Methode ist daraufhin optimiert, dem Benutzer eine möglichst breite Kenntnis der Lerngegenstände des aktuellen Satzes zu vermitteln. Durch Benutzung eines guten Gedächtnismodells wird hierbei eine effektive Lernstrategie erzielt.

[0041] Zu dem geschilderten Abfrage-Dialog sind eine Anzahl von Abwandlungen und Erweiterungen denkbar. So

kann es bspw. zu einer Frage (Definition) mehrere richtige Antworten (Vokabeln) geben. Dies kann berücksichtigt werden, bspw. indem die gespeicherten Relevanz-Maßzahlen herangezogen werden und so die relevanteren (häufigeren) Worte stärker betont werden. Die jeweiligen Sätze von Lerngegenständen können bspw. mehrere tausend Worte umfassen. Hierbei kann es sich bspw. um Lern-Gegenstände, d. h. Fachvokabular für eine bestimmte Benutzung, bspw. im Bereich der Literatur, Geschäftsleben, Technik etc. handeln.

[0042] Die Erfindung lässt sich dahingehend zusammenfassen, dass eine Vorrichtung mit Mitteln zur Aufnahme und Erkennung von Sprachsignalen und ein Verfahren zur Kommunikation mit einem elektrischen Gerät vorgestellt werden. Die Vorrichtung weist ein Verkörperungselement auf, das mechanisch bewegt werden kann. Die Position eines Benutzers wird ermittelt und das Verkörperungselement, das bspw. die Darstellung eines menschlichen Gesichts umfassen kann, wird so bewegt, dass seine Vorderseite in Richtung der Position des Benutzers weist. Am Verkörperungselement können Mikrofone, Lautsprecher und/oder eine Kamera vorgesehen sein. Der Benutzer kann einen Sprach-Dialog mit der Vorrichtung führen, wobei sich ihm das Gerät in Form des Verkörperungselements darstellt. Entsprechend der Spracheingaben des Benutzers kann ein elektrisches Gerät angesteuert werden. Ebenso ist ein Dialog des Benutzers mit dem Verkörperungselement möglich mit dem Ziel, den Benutzer zu unterrichten.

Patentansprüche

1. Vorrichtung mit Mitteln zur Aufnahme und Erkennung von Sprachsignalen (**30, 32**) und einem Verkörperungselement (**14**) mit einer Vorderseite (**44**) und Bewegungsmitteln (**24**) zur mechanischen Bewegung des Verkörperungselements (**14**), wobei Mittel (**38**) zur Bestimmung der Position eines Benutzers vorhanden sind, und die Bewegungsmittel (**24**) so angesteuert werden, dass die Vorderseite (**44**) des Verkörperungselements (**14**) in Richtung der Position des Benutzers weist.
2. Vorrichtung nach Anspruch 1, bei der Mittel (**34, 36, 18**) zur Ausgabe von Sprachsignalen vorgesehen sind.
3. Vorrichtung nach einem der vorangehenden Ansprüche, bei der das Verkörperungselement (**14**) eine anthropomorphe Darstellung, insbesondere eine Darstellung eines menschlichen Gesichts umfasst.
4. Vorrichtung nach einem der vorangehenden Ansprüche, bei der eine Anzahl von Mikrofonen (**16**) und/oder mindestens eine Kamera (**20**) vorgesehen sind, wobei bevorzugt die Mikrofone (**16**) und/oder die Kamera (**20**) am Verkörperungselement (**14**) angebracht sind.
5. Vorrichtung nach einem der vorangehenden Ansprüche, bei der Mittel zur Identifikation mindestens eines Benutzers vorgesehen sind.
6. Vorrichtung nach einem der vorangehenden Ansprüche, bei der die Bewegungsmittel (**24**) ein Schwenken des Verkörperungselements (**14**) um mindestens eine Schwenkachse ermöglichen.
7. Vorrichtung nach einem der vorangehenden An-

- sprüche, bei der mindestens ein externes elektrisches Gerät (**12**) vorgesehen ist, das durch die Sprachsignale gesteuert wird.
8. Vorrichtung nach einem der vorangehenden Ansprüche, bei der mindestens ein Lautsprecher (**18**) zur Ausgabe von akustischen Signalen, und mindestens ein Mikrofon (**16**) zur Aufnahme von akustischen Signalen vorgesehen ist, wobei eine Signalverarbeitungseinheit (**30**) zur Verarbeitung der aufgenommenen akustischen Signale vorgesehen ist, wobei Anteile des Signals, die auf vom Lautsprecher (**18**) ausgegebene akustische Signale zurückgehen, unterdrückt werden.
 9. Vorrichtung nach einem der vorangehenden Ansprüche, bei der Mittel zur Führung eines Dialogs zur Unterrichtung eines Benutzers vorgesehen sind, wobei in dem Dialog dem Benutzer visuell und/oder per Audio Aufgaben gestellt werden, und die Antworten des Benutzers mittels einer Tastatur und/oder durch ein Mikrofon aufgenommen werden.
 10. Vorrichtung nach Anspruch 9, bei der die Dialogmittel Speichermittel für einen Satz von Lerngegenständen umfassen, wobei für jeden Lerngegenstand mindestens eine Aufgabe, eine Lösung und eine Maßzahl für die Dauer seit einer Bearbeitung der Aufgabe durch den Benutzer gespeichert sind, und wobei die Dialogmittel so ausgebildet sind, dass Lerngegenstände ausgewählt und abgefragt werden, indem die Aufgabe dem Benutzer gestellt und die Antwort des Benutzers mit der gespeicherten Lösung verglichen werden, wobei bei der Auswahl der Lerngegenstände die gespeicherte Maßzahl berücksichtigt wird.
 11. Verfahren zur Kommunikation eines Benutzers mit einem elektrischen Gerät (**12**), bei dem die Position eines Benutzers ermittelt wird, und ein Verkörperungselement (**14**) so bewegt wird, dass eine Vorderseite (**44**) des Verkörperungselements (**14**) in Richtung des Benutzers weist, und Sprachsignale des Benutzers aufgenommen und verarbeitet werden.
 12. Verfahren nach Anspruch 11, bei dem entsprechend der aufgenommenen Sprachsignale das elektrische Gerät (**12**) gesteuert wird.

Hierzu 2 Seite(n) Zeichnungen

- Leerseite -

Fig. 1

